



Free-form image registration of human cochlear CT data using skeleton similarity as anatomical prior

Kjer, Hans Martin; Fagertun, Jens; Vera, Sergio; Gil, Debora; González Ballester, Miguel Angel; Paulsen, Rasmus Reinhold

Published in:
Pattern Recognition Letters

Link to article, DOI:
[10.1016/j.patrec.2015.07.017](https://doi.org/10.1016/j.patrec.2015.07.017)

Publication date:
2016

Document Version
Publisher's PDF, also known as Version of record

[Link back to DTU Orbit](#)

Citation (APA):
Kjer, H. M., Fagertun, J., Vera, S., Gil, D., González Ballester, M. A., & Paulsen, R. R. (2016). Free-form image registration of human cochlear CT data using skeleton similarity as anatomical prior. *Pattern Recognition Letters*, 76(June), 76-82. <https://doi.org/10.1016/j.patrec.2015.07.017>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.



Free-form image registration of human cochlear μ CT data using skeleton similarity as anatomical prior[☆]



H. Martin Kjer^{a,*}, Jens Fagertun^a, Sergio Vera^{b,c}, Debora Gil^{c,f}, Miguel Ángel González Ballester^{d,e}, Rasmus R. Paulsen^a

^a Department of Applied Mathematics and Computer Science, Technical University of Denmark, Copenhagen, Denmark

^b Alma IT Systems, Barcelona, Spain

^c Computer Vision Center, Universitat Autònoma de Barcelona, Barcelona, Spain

^d Department of Information and Communication Technologies, Universitat Pompeu Fabra, Barcelona, Spain

^e Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain

^f Serra Hùnter fellow

ARTICLE INFO

Article history:

Available online 30 July 2015

Keywords:

Image registration

Regularization

Skeleton

Cochlea

Inner ear

ABSTRACT

Better understanding of the anatomical variability of the human cochlear is important for the design and function of Cochlear Implants. Proper non-rigid alignment of high-resolution cochlear μ CT data is a challenge for the typical cubic B-spline registration model. In this paper we study one way of incorporating skeleton-based similarity as an anatomical registration prior. We extract a centerline skeleton of the cochlear spiral, and generate corresponding parametric pseudo-landmarks between samples. These correspondences are included in the cost function of a typical cubic B-spline registration model to provide a more global guidance of the alignment. The resulting registrations are evaluated using different metrics for accuracy and model behavior, and compared to the results of a registration without the prior.

© 2015 The Authors. Published by Elsevier B.V.

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Image registration and establishment of correspondences between data is a common challenge in biomedical image analysis. The best registration model is largely case-dependent, influenced by the anatomy, the involved imaging modalities, the desired end-goal, etc. [11,17,18]. In cases that require large and complex deformations finding the optimal registration procedure becomes a challenging task. As the amount of parameters in the transformation model increases it becomes more and more difficult for the optimization to avoid local minima. In these cases, it is often required to include some additional prior knowledge or regularization/constraints to efficiently solve the registration.

The challenging case presented in this paper is the task of registering data of the (human) inner ear. This structure controls the sensation of hearing and balance, and an understanding of the anatomy and anatomical variability plays an important part in utilizing the full potential of Cochlear Implants [30]. Detailed anatomical

models have interesting patient-specific applications as they can provide information about the type of electrode design that suits the anatomy of the user [29], or by allowing improvements to the implant programming based on simulations mimicking the actual anatomical and physiological situation [6].

The anatomy of the inner ear is composed of the vestibular system and the cochlea. Image registration of the latter is challenging for a couple of reasons, and will be the focus of this work. The human cochlea is a spiral structure with outer dimensions of approximately $10 \times 8 \times 4$ mm. The size and the shape of the spiral can vary extensively. On average, the cochlea winds 2.6 turns [9] but can approach up to three full turns – corresponding to a difference in the order of 1–2 mm following the path of the spiral. The separation between the cochlear turns is typically one order of magnitude smaller. Specially deformations to properly align the most apical region of spiral have been difficult to model to our experience. Further, the whole spiral is a tube-like structure (see Fig. 1, right) with a large degree of self-similarity in the cross-sections. This lack of distinct features makes it difficult to identify corresponding anatomical positions across samples.

The desired registration model should not just expand or compress the apical part of the spiral to align two samples, but rather

[☆] This paper has been recommended for acceptance by Punam Kumar Saha.

* Corresponding author. Tel.: +4545255228.

E-mail address: hmkj@dtu.dk (H.M. Kjer).

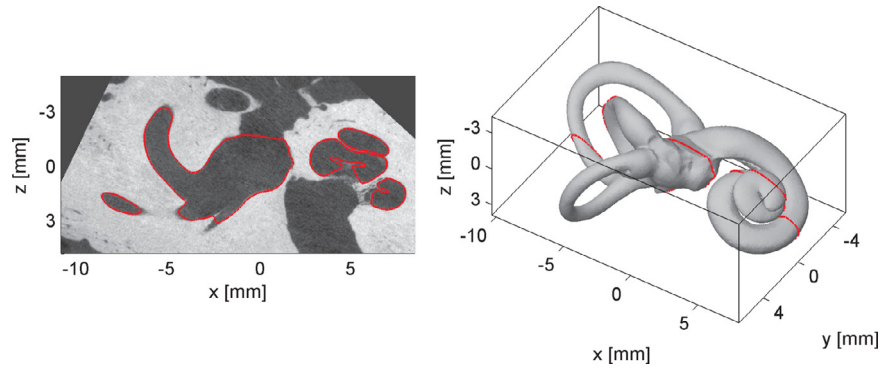


Fig. 1. Left: Impression of the μ CT data and segmentation (red border). Notice the small spacing separating the cochlear turns (right side of CT image), the weak contrast towards internal cochlea borders, and the opening into the middle ear cavity (middle of the image). Right: The corresponding surface model provides an overview of the inner ear topology. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

model a change along the entire spiral. Essentially the model should be able to handle very local deformations while still adhering to the global structure of the samples. This type of behavior is usually not native to non-rigid registration models without some kind of prior or regularization included.

Modifications to a registration model to include such prior knowledge have been studied previously. A way of introducing anatomical shape priors is the use of a statistical shape model [4,10]. However, building statistical shape models is in itself a labor intensive task rivaling if not surpassing the task of the registration, as the prerequisite for building the model is data that is already registered to have correspondences.

A multitude of physical constraints have also been proposed as regularizations. For example, local tissue rigidity can be enforced in specified areas [23], or conditions of incompressibility or volume-preservation can be applied [20]. However, finding the suitable physical constraint for a registration task is not straightforward, as this is case- and application dependent.

In the work of [3] an articulated skeleton model was pre-registered to intra-mouse data studies in order to recover large pose-differences between data acquisitions. The presented application is narrow in its scope, but the registration methodology of using landmark correspondences as regularization is more generally applicable, thus we adopt this approach for this work.

In this paper we explore the potential of using the skeleton of a surface object as an anatomical prior in free-form registrations using a B-spline transformation model.

Skeletonization of a volume or surface is a research field in itself [7,22]. The skeleton provides a global description of shape in a simplified and structured form. Matching based on skeleton similarity could provide a global anatomical guidance or regularization to a locally defined free-form image registration procedure with a high resistance to noise compared to using only the image intensity similarity. The use of skeleton similarity in image registrations should be applicable to many different problems and there are many published methods and approaches for finding and matching the skeletons for differing types of data and geometries [24,26]. Skeleton correspondence has been seen in image registration tasks before, relating to for instance 2D/3D multi-modal registration [15] and matching of vessels in time-series angiography data [27]. More related to our approach is the work of [25], where multiple different shape features were calculated from surface objects and transformed into vector-valued 2D feature images, which were aligned with a classic image registration formulation. Skeleton features were used for global alignment in the coarser levels of the registration. Our strategy is similar although the prior will be included into the registration model differently.

The purpose of this study is to test and evaluate deformable registration using a B-spline transformation model on a series of inner

ear datasets with/without the use of skeleton-based similarity in the registration model.

The rest of the paper is structured as follows. Section 2.1 provides a description of the data and the processing, and Section 2.2 contains the procedure for finding skeletons and their similarity across datasets. The registration models and their evaluation are detailed in Sections 2.3 and 2.4 respectively. The results are presented in Section 3 and discussed in Section 4.

2. Material and methods

2.1. Data and processing

A collection of 17 dried temporal bones from the University of Bern were prepared and scanned with a Scanco Medical μ CT100 system. The data was reconstructed and processed to obtain image volumes of 24 micron isotropic voxel-sizes containing the inner ear (Fig. 1, left).

Image segmentation: The border of the inner ear was segmented in all datasets semi-automatically using ITK-SNAP [31]. On standard CT images the cochlear will appear to have a circular cross-section. Due to the higher resolution of μ CT and the sample preparation it becomes possible to see the *lamina spiralis*. It is a bony ridge structure that traverses the entire cochlea from the spiral central direction, partially separating the cochlear into two chambers (*scala tympani* and *scala vestibuli*) and creating a 'U'-shaped cross-section (see Fig. 1, left). The semi-automatic tool in the segmentation software was critical for achieving smooth and rounded segmentations in data with that kind of resolution, and for reducing the amount of manual work. The images contain some openings, less well-defined regions and non-anatomical artifacts that had to be manually handled to obtain comparable segmentations across datasets. For this reason a segmentation of one dataset easily amounts to 12–15 h, but it is a requirement for having a ground truth and a correct representation of the object. A surface model was generated for each dataset using Marching Cubes [16] followed by a surface reconstruction [19] to obtain a well-formed triangular mesh (Fig. 1, right).

2.2. Skeletonization

Implementing and comparing skeletonization methodologies is not the scope of this work. Our aim is to find a simple method for generating skeleton correspondence between samples that can easily be included in a registration model thus allowing us to explore the potential of the approach. The object topology is an important consideration when working with skeletons, as this poses a restriction for certain methods. To avoid working with a genus 3 surface, we exclude the vestibular system and focus only on a skeleton of the spiral shaped cochlea.

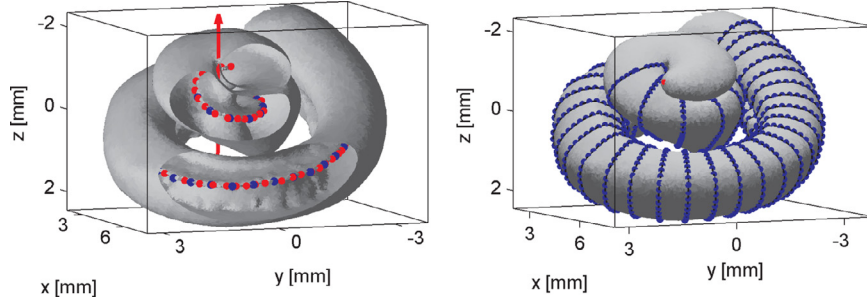


Fig. 2. The cochlear skeletonization. Red annotations are manually determined information (Left: z_i^{Man} and central axis, Right: cochlear apex). Blue annotations represent parametric pseudo-landmarks (Left: z_i^S , Right: z_i^{LM}). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

Initial experimentation showed that methods for automatically finding the curved skeleton [7], medial curve or centerlines [5] tended to be attracted to the larger of the two scalae and thus creating inconsistent skeletons across samples. Medial surface representations [12,28] were also found challenging to generate with desirable topological consistency and smoothness. To keep things simple we propose to use a set of corresponding pseudo-landmarks, z_i^{LM} , of the cochlear surfaces obtained from a parametric ‘curved skeleton’ following the *lamina spiralis* ridge.

We manually define the following information (marked red on Fig. 2) in each dataset to determine our corresponding pseudo-landmarks: The cochlear apex landmark (A_i), the central axis of the cochlear spiral (defined by a unit direction vector, \vec{n}_i , and an axis point, C_i), and a set of points (approx. 100 per dataset) following the ridge line of the *lamina spiralis* from the beginning of the cochlear base to the end in the apical turn (z_i^{Man}).

From the above information, we generate a naive parametric model of the cochlea. First, we create a parametric description of the cochlea skeleton using z_i^{Man} , from which we determine 37 corresponding positions on the skeleton with equal arc-length (z_i^S). Secondly, we extract planar surface cross-section at each of the points, p , in z_i^S . The cross-section plane is determined by the tangent of the skeleton at p . Each cross-section of the surface mesh is then parameterized using 40 points (\vec{n}_i provides the reference for orientation and starting point in each cross-section parameterization). These cross-sectional points together with the apex landmark provides a set, z_i^{LM} , of 1481 corresponding surface pseudo-landmarks (Fig. 2, right) to be included in a registration model. Finding the cochlea cross-section in the apical region of the cochlear can potentially lead to some ambiguity, as they could intersect with themselves. To avoid this the skeleton cross-sections in the apical turn were not included.

2.3. Image registration

The registration procedure follows a common work-flow. One dataset was chosen as the reference, to which the remaining moving datasets were registered in two steps - rigid initialization followed by the deformable registration, both detailed in the following subsections.

2.3.1. Initial rigid alignment

There are many approaches for finding rigid transformations. The chosen procedure is independent from the skeleton information and is the same no matter the chosen deformable registration model. In that way, later comparisons of registration results are not affected by the initialization. The whole initialization procedure relies solely upon the extracted surface meshes, but the calculated rigid transformations were also applied to the gray-scale volumes and their segmentations.

Translation: Let $p_{(i,j)}$ be the j th vertex position of dataset i . A translation was applied so that the center of mass is placed in po-

sition (0,0,0), i.e. the mean vertex position, \bar{p}_i , was subtracted from all vertices. This places all datasets in a coordinate system where the inner ear center of mass of each dataset is in the origin.

Rotation: Let Σ_i be the 3×3 covariance matrix of the mesh vertex positions of dataset i (after the translation). The eigenvectors, W_i , of Σ_i provides a rotation matrix, which when applied transforms the data to the principal component directions. This essentially corresponds to fitting an ellipsoid to the point cloud and aligning the axes.

Check directions: This alignment procedure is robust due to the asymmetry of the inner ear shape (Fig. 1, right). However, the sign of a principal direction in the i th dataset could potentially be opposite compared to that of the reference. To handle this we make a simple check. The bounding box of the reference and of the moving point cloud is divided into a coarse grid. We use the sum of squared grid vertex-density difference between the two as a check metric. If the axis-flip would result in a lower metric, then the flip is made to the moving dataset. While there is no guarantee for this to work in all cases, it has worked well for our data. In principle, any kind of rigid alignment could be used instead of the one suggested here.

2.3.2. Deformable registration

The non-rigid image registration follows the formulation and framework of *elastix* [14].

The registration is done between the segmentations rather than the gray-scale volumes for two reasons. First, the μ CT data contain smaller artifacts and certain weakly contrasted edges, that were handled during the segmentation. Secondly, the registration should not be influenced by the anatomical differences in the surrounding bone structure.

The registration of the moving dataset, I_M , towards the reference, I_F , is formulated as a (parametric) transformation, T_μ , where the vector μ containing the p -parameters of the transformation model are found as an optimization of a cost function, \mathcal{C} .

$$\hat{\mu} = \arg \min_{\mu} \mathcal{C}(T_\mu, I_F, I_M) \quad (1)$$

The transformation model used in this paper is the cubic B-spline in a multi-resolution setting. We apply image smoothing with a Gaussian kernel to both the fixed and moving image. For each level of resolution the spacing between grid points and the width of the smoothing kernel follows a decreasing scheme, starting with a coarse registration that is gradually refined. The following scheme was chosen by experimentation:

Control point grid spacing (isotropic, voxels):

[144, 72, 48, 48, 36, 24, 18, 12, 6]

Width of Gaussian kernel (isotropic, voxels):

[10, 10, 1, 1, 1, 1, 1, 1, 1]

The width of the kernel was deliberately kept narrow in most levels to avoid that small and sharp features would be blurred out (for instance the separation of the cochlear turns).

The cost function used in this ‘basic’ registration set-up:

$$C_1 = \alpha \cdot S_{\text{Sim}}(\mu, I_F, I_M) + (1 - \alpha) \cdot \mathcal{P}_{\text{BE}}(\mu) \quad (2)$$

where α is a weight parameter in the interval [0,1]. The similarity term, S_{Sim} , is chosen as the sum of squared differences (SSD). The term \mathcal{P}_{BE} is the energy bending regularization used to penalize strong changes and foldings in the transformation [21]. The weighting of the similarity term was chosen to 0.9 by experimentation. Increasing α would provide more freedom for deformation of the shapes, but also increase the risk of having non-plausible anatomical results.

The optimization is solved using Adaptive Stochastic Gradient Descent [13]. The maximum number iterations was set to 2500. To reduce the computational burden of the optimization only a subset voxels are sampled for the evaluation. For each iteration 2^{14} random coordinate points were sampled. These settings were fixed for all resolutions.

2.3.3. Deformable registration with guidance from skeleton

The free-form registration set-up remains largely the same when a skeleton is included in order to make comparisons fair. The cost function is modified to include a landmark similarity term [3]:

$$C_2 = \alpha \cdot S_{\text{Sim}}(\mu, I_F, I_M) + \beta \cdot S_{\text{CP}}(\mu, \mathcal{Z}_F, \mathcal{Z}_M) + (1 - \alpha - \beta) \cdot \mathcal{P}_{\text{BE}}(\mu) \quad (3)$$

where α and β are weightings in the interval [0,1] and fulfilling $\alpha + \beta \leq 1$. The landmark similarity term, $S_{\text{CP}}(\mu, \mathcal{Z}_F, \mathcal{Z}_M)$, uses the Euclidean distance between the set of corresponding landmarks, \mathcal{Z}_F and \mathcal{Z}_M . In this way intensity-based image registration is guided with features extracted from the anatomical skeleton (i.e. using $\mathcal{Z}_i^{\text{LM}}$ from Section 2.2). By experimentation the weightings were set to $\alpha = 0.8$ and $\beta = 0.11$. The landmark similarity is kept small in order not to force the alignment, and the ratio between image similarity and bending energy regularization is kept similar to the previous set-up C_1 (Eq. (2)). Settings for the transformation model and optimizer were unchanged from the previous registration model.

2.4. Evaluation

We are interested in comparing the 16 registration results of model 1 (Eq. (2)) and model 2 (Eq. (3)) using a number of different image and mesh based metrics.

Image based evaluation: Let $I_i(\mu)$ be the moving segmentation volume after application of the resulting transformation. We compare the Dice Score [8] to the segmentation of the reference dataset, I_{Ref} .

$$\text{DSC} = \frac{2 \cdot |I_{\text{Ref}} \cap I_i(\mu)|}{|I_{\text{Ref}}| + |I_i(\mu)|} \quad (4)$$

Mesh based evaluation: We define the surface based scores as follows. Let $S_{\text{Ref}}(\mu)$ be the reference surface mesh after application of the resulting transformation. There is no direct point correspondence between the reference and the ground truth surfaces, S_i , and they each contain a varying number of vertices. Metrics are therefore based on the closest points, i.e. the minimum Euclidean distance from a point, p , to any of the points, q , in the other surface, S :

$$d(p, S) = \min_{\forall q \in S} (||p - q||_2) \quad (5)$$

The mean surface error, d_s , of each sample is defined as the average of all the closest point distances:

$$d_s = \frac{1}{N_{\text{Ref}} + N_i} \left(\sum_{\forall p \in S_{\text{Ref}}(\mu)} d(p, S_i) + \sum_{\forall p \in S_i} d(p, S_{\text{Ref}}(\mu)) \right) \quad (6)$$

where N_{Ref} and N_i are the total number of points in the reference and the moving surface respectively.

The Hausdorff distance, d_H , is the maximum of all the closest point distances:

$$d_H = \max \left\{ \max_{\forall p \in S_{\text{Ref}}(\mu)} d(p, S_i), \max_{\forall p \in S_i} d(p, S_{\text{Ref}}(\mu)) \right\} \quad (7)$$

The above mentioned metrics are very generic and will hardly be able to reflect and evaluate the change in the registration model that we intend to explore. We therefore include two additional scores, apex error and torque.

First, we calculate the Euclidean distance between apexes of the target data and of the reference.

$$d_A = ||A'_{\text{Ref}}(\mu) - A_i||_2 \quad (8)$$

The apex is one of the few locations on the cochlea that can be placed relatively precisely. Even though an arc-length distance might be more correct to use, the Euclidean apex error should be indicative of the registration model behavior in the apical region, even though this point is also included in the registration model.

Secondly, we look at the differences in the vector deformation fields obtained by the registration models. The cochlear samples have a different number of turns, and we wish to evaluate the registration models ability to capture this rotational behavior of the anatomy. Our postulation and assumption is that this ability of the registration model should correlate with the ‘torque’, τ , on the central axis of the cochlear exerted by the deformation field.

Let the force vector, \vec{F}_p , on the vertex, p , in the reference mesh be defined simply as the vector between the vertex position before and after application of the registration transformation:

$$\vec{F}_p = p(\mu) - p$$

Further, we can calculate the perpendicular arm from the central axis to the mesh vertex, \hat{v}_p . This vector is normalized to unit length, so that the vertices farther from the axis will not contribute with a greater torque.

The scalar projection of the force vector, F_p , onto the unit arm that is perpendicular to both the central axis and \hat{v}_p is then the acting force contributing to the torque:

$$F_p = \vec{F}_p \cdot (\vec{n} \times \hat{v}_p)$$

Using this local vertex torque force leads to our definition of the torque of the registration:

$$\tau = \frac{1}{N_{\text{Ref}}} \sum_{\forall p \in S_{\text{Ref}}} F_p = \frac{1}{N_{\text{Ref}}} \sum_{\forall p \in S_{\text{Ref}}} (p(\mu) - p) \cdot (\vec{n} \times \hat{v}_p) \quad (9)$$

3. Results

The registrations were done on a desktop with a quad-core 3.6 GHz processor, 64 GB RAM, running elastix v4.7. The average time per registration was approximately 0.8 h and we observed no notable

Table 1

Statistics of registration evaluation metrics, reported as the mean \pm 1 std. Model 1 is the non-rigid image registration model and model 2 the non-rigid image registration model incorporating a skeleton prior.

Metric	Dice score	Surface error [mm]	Hausdorff [mm]	Apex error [mm]	Avg. torque [mm ²]
Model 1	0.96 \pm 0.01	0.040 \pm 0.01	0.69 \pm 0.24	1.01 \pm 0.59	−0.04 \pm 0.09
Model 2	0.95 \pm 0.01	0.045 \pm 0.01	0.73 \pm 0.35	0.69 \pm 0.52	−0.53 \pm 0.28

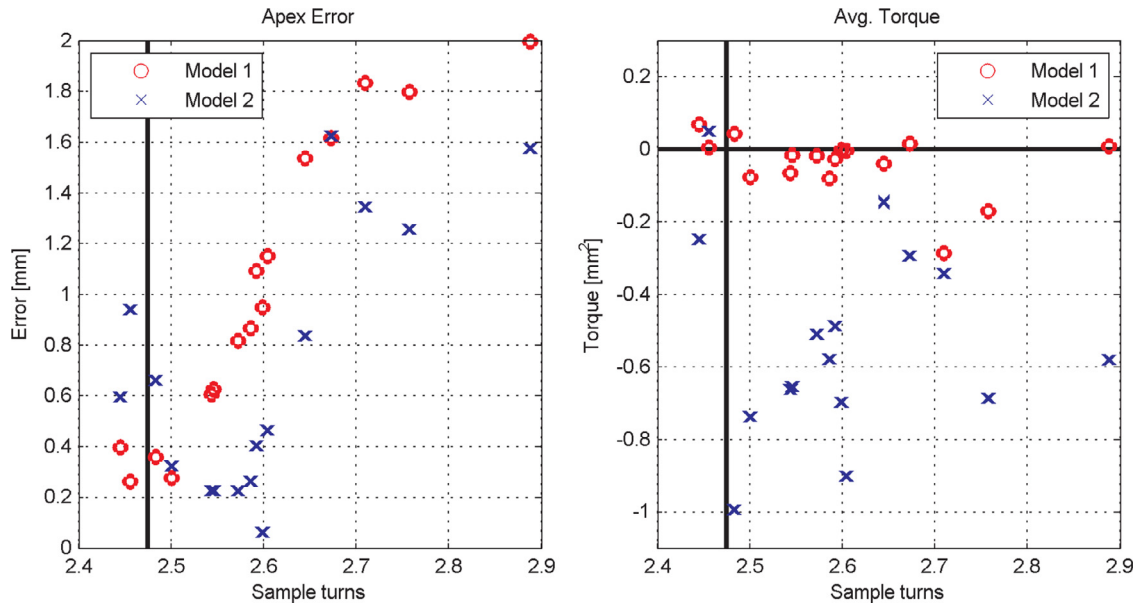


Fig. 3. Sample-wise apex error (Left) and average torque (Right) plotted against the number of cochlear turns of the target samples. Vertical black line indicate the number of turns in the reference sample.

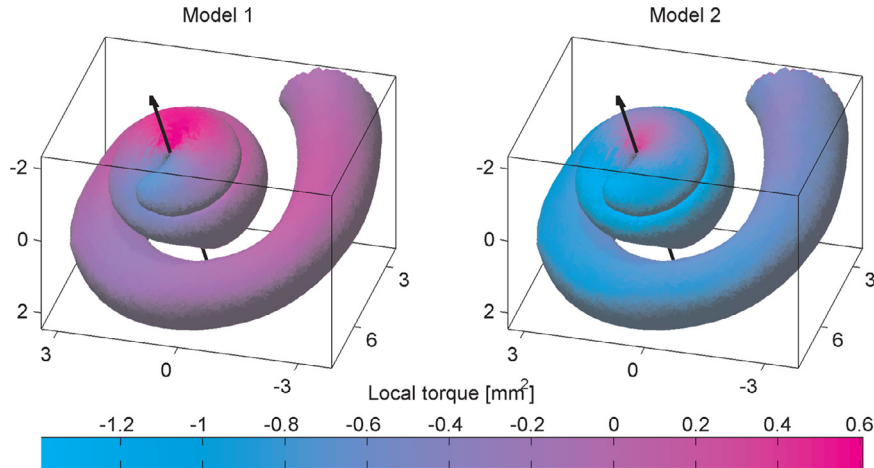


Fig. 4. Qualitative difference in the local torque acting on the cochlea central axis (black vector). The target sample has 2.60 turns, compared to the 2.46 of the reference (the shown surface). Positive direction of the central axis is defined from the cochlea base towards the apex.

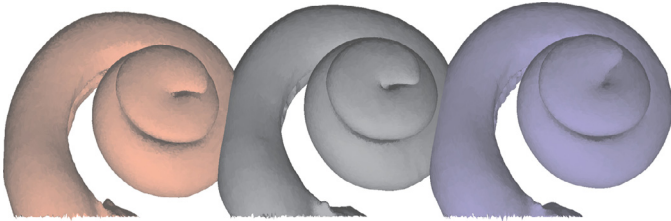


Fig. 5. The visual difference between registration models. The reference surface is deformed using either model 1 (purple, right) or model 2 (red, left) to align with the target sample (grey, middle). The surfaces have been moved apart to avoid overlap between shapes.

difference in run times or convergence speed between the two registration models.

The statistics of the different metric scores are presented in Table 1. Fig. 3 elaborates on the sample-wise apex error and torque metric, and Figs. 4 and 5 show the qualitative difference between the registration models.

The general metrics (DSC, d_s , d_H) show a small decrease in performance accuracy for model 2.

From Fig. 3 it is observed that the apex errors of model 1 grow more or less proportionally to the discrepancy in cochlear turns. The torque is close to zero on average. These observations reflect that model 1 only adapts very locally and behaves indifferently with regards to the turning of the target shape. I.e. the resulting cochlear shapes after registration have little variation in the turns.

The apex errors are seen to be generally lower for model 2. Note, that the apex landmark used to calculate this error was a part of the optimization procedure. That the error is reduced is therefore no surprise and it is a biased metric for considering the model accuracy and precision. However, it provides a summarizing pseudo-measure of how much more turning registration model 2 on average is able to capture, which is further illustrated in Fig. 5. For very large differences in cochlear turns it would seem that both of the registration models have trouble with aligning the apices.

The torque of model 2 is in most of the cases negative. This indicate vector fields pointing more tangentially in the direction of the spiral towards to the apical region. This would be the expectation as most of the target samples have more turns than the reference. The torque is not a measure of accuracy nor precision. The torque merely

provides a simple quantification of the overall rotation of the cochlear shape. Further it gives a good way of illustrating the differences between the registration models as demonstrated in Fig. 4.

4. Discussion

The Dice Score, surface error and Hausdorff distance serves as very general metrics for evaluating the local adaptability of the registration models. Further, they indicate the general accuracy and precision that we are achieving with the data. The performance with model 2 was decreased on these scores. It would seem that we are trading some local adaption for guiding the model with the landmarks. The determination of the skeleton inherently carries some uncertainties. By introducing the landmarks into the registration model extra noise is added to the procedure. It may happen that a poor skeleton estimate is drawing the spiral in the wrong direction. By providing a more robustly determined skeleton that additionally could fully reach the most apical turn, we expect that the performance of model of 2 could be increased.

The current approach by basing the skeletons on manually picked medial points is only feasible because of the low number of data samples, and because of the specific anatomy where the ridge of the *lamina spiralis* is easily identifiable and happens to define a medial line of the object. Manually placing medial points in a 3D model would generally be impossible.

Aspects of the skeletonization and its influence can be studied furthermore. For instance the number of landmarks used to represent the skeleton. By experimentation we found an amount of cross-sections that appeared to work, but the number of landmarks per cross-section could potentially be reduced. However, the primary concern is the current lack of information in the most apical cochlear turn. For this to be included it would be interesting to look into other skeletal representations. That would in turn potentially require a different way of measuring the similarity of skeletons and possibly an extension to the registration framework to accommodate this. It holds an interesting research potential as both the field of skeletonization and image registration are well-researched areas, but so far joining the two have received little focus. A reason might be the challenge in automatically obtaining consistent skeletons from volumetric data. In this work the skeletons were based on the surface models (i.e. the data segmentation), which in many cases are difficult and/or time-consuming to obtain. Ideally the skeletons should be extracted from volumetric gray-scale data similar to the work of [1,2].

Using the B-spline grid as the transformation model in the registration has limitations. Choosing a fluid- or optical flow-based model [18] could potentially be more suited for this kind of spiral anatomy. Alternatively, the performance of the B-spline approach could perhaps be improved with some data preprocessing. If the cochlea was unfolded, possibly based on the skeleton cross-sections, it would be in a space more suited for a B-spline grid transformation. Along the same line of thinking, the deformation control points could be placed in a non-cubic grid structure favoring the spiral nature of the data. However, these suggestions may be difficult to realize and involves adapting the registration method to one very specific task or anatomy. In this and potentially other cases finding a skeleton and including it into the a registration model may be an easier or more feasible approach. The results reflect that it is possible to modify and regularize the registration by using skeleton similarity as a prior, even though there is room for improvements in our methodology.

The registration parameters used in this work were manually determined. A set of parameters that works well on all data samples while running within a reasonable time frame can be difficult to find. Regarding the choice of metric weights, an interval of $\alpha = 0.7 - 0.9$, would seem to be the most appropriate for model 1. Higher α increases the flexibility of the model, which is needed for capturing the cochlear turning. However, increasing beyond 0.9 made some

cases fail. In particular the behavior of the deformations in the semi-circular canals performed poorly. The same holds true for model 2. For having a fair comparison between the registration models, the same relative weight of the image similarity and bending energy metric was kept. Having $\beta < 0.15$ was found to be reasonable. Forcing more weight on the landmarks could result in too strong deformations in some cases, and going much lower counters the idea of having the landmarks. Variable metric weights throughout the resolutions were also tested for model 2. I.e. a scheme where a strong weighting was placed on the landmarks in the initial resolutions and then gradually reduced. It worked well in some cases only, so to keep the registration models comparable the fixed weightings scheme was used. Regarding the optimization only the default optimizer and automatically determined settings were used. A number of samples in the range of $2^{14} - 2^{17}$ and a maximum number of iterations between 1000 and 2500 seemed to produce stable results. Tweaking of registration parameters could result in minor changes of the performance scores, but the same tendencies of the registration models would be observed.

The local torque forces (Fig. 4) provides the most qualitative view of the differences between the registration models. There is no ground truth torque, but it illustrates that the normal registration model is very local in its adaption, whereas model 2 provides more turning in the region where the skeleton is defined. Ideally we could have shown a more convincingly stronger negative correlation (Fig. 3) between the differences in the cochlear turns and the average size of the torque. However, we have a low number of samples and the registration also has to deal with general differences in the size and orientation of the samples apart from the turning. In future work the torque could perhaps even be used as a regularization in the registration model, where it could favor a constant torque in the B-spline grid points near the spiral.

5. Conclusion

We have demonstrated the use of parameterized skeletons to act as an anatomical similarity prior for cubic B-spline non-rigid image registration of cochlea μ CT data.

The B-spline transformation model is only locally defined, and the registration is challenged by the spiral nature of the cochlear data, that is locally similar throughout all turns. We have created a simple parameterized skeleton, and included corresponding parametric pseudo-landmarks into the registration cost function. The skeleton provides global similarity to the registration model, that allows cochlear turning to be captured to a larger degree.

Developments in the use of skeletons in the non-rigid alignment could lead to better registration models. This requires further work in combining research in skeletonization with the field of image registration.

Acknowledgments

HEAR-EU is a collaborative project between Alma Medical Imaging, Med-El, Scanco Medical AG, the Univ. of Bern, Univ. Pompeu Fabra and the Tech. Univ. of Denmark. We would like to thank Scanco and the University/University Hospital of Bern for acquiring and providing the data.

The research leading to HEAR-EU results has received funding from the [European Union Seventh Frame Programme \(FP7/2007-2013\)](#) under grant agreement n 304857.

Debora Gil is a Serra Húnter fellow and supported by Spanish Government Project TIN2012-33116.

References

- [1] S.S. Abeyasinghe, M. Baker, W. Chiu, T. Ju, Segmentation-free skeletonization of grayscale volumes for shape understanding, in: *Proceedings of IEEE International Conference on Shape Modeling and Applications* 2008, 2008, pp. 63–71.

- [2] E. Antúnez, L. Guibas, Robust extraction of 1D skeletons from grayscale 3D images, in: *Proceedings of the International Conference on Pattern Recognition*, 2008.
- [3] M. Baiker, M. Staring, C.W.G.M. Löwik, J.H.C. Reiber, B.P.F. Lelieveldt, Automated registration of whole-body follow-up microCT data of mice, *Med. Image Comput. Comput. Assist. Interv.* 14 (Pt 2) (2011) 516–523, doi:[10.1007/978-3-642-23629-7_63](https://doi.org/10.1007/978-3-642-23629-7_63).
- [4] F.F. Berendsen, U.A. van der Heide, T.R. Langerak, A.N.T.J. Kotte, J.P.W. Pluim, Free-form image registration regularized by a statistical shape model: application to organ segmentation in cervical MR, *Comput. Vision Image Underst.* 117 (9) (2013) 1119–1127, doi:[10.1016/j.cviu.2012.12.006](https://doi.org/10.1016/j.cviu.2012.12.006).
- [5] S. Bouix, K. Siddiqi, A. Tannenbaum, Flux driven automatic centerline extraction, *Med. Image Anal.* 9 (3) (2005) 209–221, doi:[10.1016/j.media.2004.06.026](https://doi.org/10.1016/j.media.2004.06.026).
- [6] M. Ceresa, N. Mangado Lopez, H. Dejea Velardo, N.C. Herrezuelo, P. Mistrik, H.M. Kjer, S. Vera, R.R. Paulsen, M.A. González Ballester, Patient-specific simulation of implant placement and function for cochlear implantation surgery planning, in: *Proceedings of MICCAI*, 2014, pp. 49–56, doi:[10.1007/978-3-319-10470-6_7](https://doi.org/10.1007/978-3-319-10470-6_7).
- [7] N.D. Cornea, D. Silver, P. Min, Curve-skeleton properties, applications, and algorithms, *IEEE Trans. Vis. Comput. Graphics* 13 (3) (2007) 530–548, doi:[10.1109/TVCG.2007.1002](https://doi.org/10.1109/TVCG.2007.1002).
- [8] L.R. Dice, Measures of the amount of ecologic association between species, *Ecology* 26 (3) (1945) 297–302.
- [9] E. Erixon, H. Högstorp, K. Wadin, H. Rask-Andersen, Variational anatomy of the human cochlea: implications for cochlear implantation, *Otol. Neurotol.* 30 (1) (2009) 14–22, doi:[10.1097/MAO.0b013e31818a08e8](https://doi.org/10.1097/MAO.0b013e31818a08e8).
- [10] T. Heimann, H.-P. Meinzer, Statistical shape models for 3D medical image segmentation: a review, *Med. Image Anal.* 13 (4) (2009) 543–563, doi:[10.1016/j.media.2009.05.004](https://doi.org/10.1016/j.media.2009.05.004).
- [11] D. Hill, P. Batchelor, M. Holden, D. Hawkes, Medical image registration, *Phys. Med. Biol.* 46 (3) (2001) R1–R45, doi:[10.1088/0031-9155/46/3/201](https://doi.org/10.1088/0031-9155/46/3/201).
- [12] T. Ju, M.L. Baker, W. Chiu, Computing a family of skeletons of volumetric models for shape description, *Comput.-Aided Des.* 39 (5) (2007) 352–360, doi:[10.1016/j.cad.2007.02.006](https://doi.org/10.1016/j.cad.2007.02.006).
- [13] S. Klein, J.P.W. Pluim, M. Staring, M.A. Viergever, Adaptive stochastic gradient descent optimisation for image registration, *Int. J. Comput. Vis.* 81 (3) (2009) 227–239, doi:[10.1007/s11263-008-0168-y](https://doi.org/10.1007/s11263-008-0168-y).
- [14] S. Klein, M. Staring, K. Murphy, M.A. Viergever, J.P. Pluim, Elastix: a toolbox for intensity-based medical image registration, *IEEE Trans. Med. Imaging* 29 (1) (2010) 196–205.
- [15] A. Liu, E. Bullitt, S. Pizer, 3D/2D registration via skeletal near projective invariance in tubular objects, in: *Proceedings of Medical Image Computing and Computer-Assisted Intervention - MICCAI'98*, 1496, 1998, pp. 952–963.
- [16] W.E. Lorensen, H.E. Cline, Marching cubes: a high resolution 3D surface construction algorithm, *SIGGRAPH Comput. Graph.* 21 (4) (1987) 163–169, doi:[10.1145/37402.37422](https://doi.org/10.1145/37402.37422).
- [17] J.B.A. Maintz, M.A. Viergever, A survey of medical image registration, *Med. Image Anal.* 2 (1) (1998) 1–36, doi:[10.1016/S1361-8415\(98\)80001-7](https://doi.org/10.1016/S1361-8415(98)80001-7).
- [18] F.P.M. Oliveira, J.M.R.S. Tavares, Medical image registration: a review, *Comput. Methods Biomech. Biomed. Eng.* 17 (2) (2014) 73–93, doi:[10.1080/10255842.2012.670855](https://doi.org/10.1080/10255842.2012.670855).
- [19] R. Paulsen, J. Baerentzen, R. Larsen, Markov random field surface reconstruction, *IEEE Trans. Visual. Comput. Graphics* 16 (4) (2010) 636–646, doi:[10.1109/TVCG.2009.208](https://doi.org/10.1109/TVCG.2009.208).
- [20] T. Rohlfing, C. Maurer, D. Bluemke, M. Jacobs, Volume-preserving nonrigid registration of MR breast images using free-form deformation with an incompressibility constraint, *IEEE Trans. Med. Imaging* 22 (6) (2003) 730–741, doi:[10.1109/TMI.2003.814791](https://doi.org/10.1109/TMI.2003.814791).
- [21] D. Rueckert, L. Sonoda, C. Hayes, D. Hill, M. Leach, D. Hawkes, Nonrigid registration using free-form deformations: application to breast mr images, *IEEE Trans. Med. Imaging* 18 (8) (1999) 712–721, doi:[10.1109/42.796284](https://doi.org/10.1109/42.796284).
- [22] K. Saeed, M. Tabeedzki, M. Rybnik, M. Adamski, K3M: a universal algorithm for image skeletonization and a review of thinning techniques, *Int. J. Appl. Math. Comput. Sci.* 20 (2) (2010) 317–335, doi:[10.2478/v10006-010-0024-4](https://doi.org/10.2478/v10006-010-0024-4).
- [23] M. Staring, S. Klein, J.P.W. Pluim, A rigidity penalty term for nonrigid registration, *Med. Phys.* 34 (11) (2007) 4098–4108, doi:[10.1118/1.2776236](https://doi.org/10.1118/1.2776236).
- [24] H. Sundar, D. Silver, N. Gagvani, S. Dickinson, Skeleton based shape matching and retrieval, in: *Proceedings of Shape Modeling International 2003 (SMI)*, 2003, pp. 130–139.
- [25] L. Tang, G. Hamarneh, Smrfi: Shape matching via registration of vector-valued feature images, in: *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2008, pp. 1–8.
- [26] J.W. Tangelde, R.C. Veltkamp, A survey of content based 3D shape retrieval methods, in: *Proceedings of Shape Modeling International SMI 2004*, 2004, pp. 145–156.
- [27] B. Tom, S. Efstratiadis, A. Katsaggelos, Motion estimation of skeletonized angiographic images using elastic registration, *IEEE Trans. Med. Imaging*, 13 (3) (1994) 450–460.
- [28] S. Vera, D. Gil, A. Borrás, M.G. Linguraru, M.Á. González Ballester, Geometric steerable medial maps, *Mach. Vis. Appl.* 24 (6) (2013) 1255–1266, doi:[10.1007/s00138-013-0490-4](https://doi.org/10.1007/s00138-013-0490-4).
- [29] S. Vera, F. Perez, C. Balust, R. Trueba, J. Rubió, R. Calvo, X. Mazaira, A. Danas- ingh, L. Barazzetti, M. Reyes, M. Ceresa, J. Fagertun, H. Kjer, R. Paulsen, M. González Ballester, Patient Specific Simulation for Planning of Cochlear Implantation Surgery, *Lecture Notes in Computer Science*, Springer, pp. 101–108. doi:[10.1007/978-3-319-13909-8_13](https://doi.org/10.1007/978-3-319-13909-8_13).
- [30] B.S. Wilson, M.F. Dorman, Cochlear implants: a remarkable past and a brilliant future, *Hear. Res.* 242 (1–2) (2008) 3–21, doi:[10.1016/j.heares.2008.06.005](https://doi.org/10.1016/j.heares.2008.06.005).
- [31] P.A. Yushkevich, J. Piven, H.C. Hazlett, R.G. Smith, S. Ho, J.C. Gee, G. Gerig, User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability, *Neuroimage* 31 (3) (2006) 1116–1128.